Theoretical Economics Letters, 2022, 12, 1773-1799
https://www.scirp.org/journal/tel
ISSN Online: 2162-2086
ISSN Print: 2162-2078

Estimation of Long-Term Profitability of Startups: An Experimental Analysis

Erkki K. Laitinen, Teija Laitinen*

School of Accounting and Finance, University of Vaasa, Vaasa, Finland
Email: erkki.k.laitinen@uwasa.fi, *teija.laitinen@uwasa.fi

How to cite this paper: Laitinen, E. K., & Laitinen, T. (2022). Estimation of Long-Term Profitability of Startups: An Experimental Analysis. *Theoretical Economics Letters, 12,* 1773-1799.
https://doi.org/10.4236/tel.2022.126096

Received: October 27, 2022
Accepted: December 26, 2022
Published: December 29, 2022

Abstract

The objective is to assess the performance of different methods to derive an estimate of internal rate of return (IRR) for startups. Koyck transformation is first used to estimate the parameters of a distributed revenue lag model which are then used to derive IRR. For estimation different scenarios of artificial time series of expenditure and revenue are constructed to describe the early years of startups. These scenarios are based on different parameter values of the distributed lag function and are classified into nine experiments. The performance of the following six different estimation methods are compared with each other in these nine experiments: unrestricted OLS, OLS through the origin (RTO), restricted OLS, Least Absolute Deviation (LAD), Ridge Regression (RR), and restricted Maximum Likelihood (ML). The experimental results indicate that the most efficient estimation method is the Ordinary Least Squares (OLS) method where the regression is forced through the origin (RTO). The least efficient method is the unrestricted OLS, which emphasizes the importance of RTO.

Keywords

Startups, Long-Term Profitability, IRR, Distributed Lag Model, Growth, Koyck Transformation, Experimental Data, Estimation Methods

1. Introduction

Startups have a remarkable impact on economies that are willing to invest in innovations and growth. The first years of startups are typically full of financial difficulties and about 45% - 50% of startups fail during the first five years, in a valley of death (U.S. Bureau of Labor Statistics, 2022; Eurostat, 2022). The importance of startups and their high failure risk have made the early years of startups a popular theme in business research but also in government policy debate

(Davila, Foster, He, & Shimizu, 2015). It is often argued that the failures are a result of that startups having insufficient finance to carry out their business in the first years (Huynh, Petrunia, & Voia, 2012). Startups have (due to a high risk to fail) difficulties in raising equity capital and are forced to rely on borrowing. Moreover, startups suffer from insufficiency of internal finance that should, however, be the main source of financial capital (Zingales, 1998). In startups, insufficiency of internal finance is strongly related to profitability, growth, and the length of lag in revenue generation. If a startup grows quickly, has a low profitability, and has a long lag in revenue generation, it evidently runs into difficulties with internal financing in the first years.

The importance of startups makes it essential for investors, financers, entrepreneurs, and public actors to understand how to assess the long-term profitability of startups. The task is difficult, since the first years can be difficult for a startup that is anyway profitable but grows too quickly or generates revenues too slowly. In these early years, all yearly financial ratios can refer to serious financial difficulties even if the long-term profitability of a startup is good (Laitinen, 1992; Laitinen, 2017). It is expected that the most efficient firms should receive financing and have higher valuations than worse firms (Ak, Dechow, Sun, & Wang, 2013). However, it is not an easy task to assess the efficiency of startups on the basis of early financial time series which include intertwined information that is difficult to interpret. The critical task is to assess whether the actual relation of revenues to expenditures (that determines the amount of internal financing), is a result of low profitability (inefficiency), fast growth, or slow generation of revenues.

When assessing the efficiency of a startup, long-term profitability as measured by the internal rate of return (IRR) plays the central role. The main problem is that IRR is not directly observable but must be estimated from non-steady and short time series of startups, which makes the use of complicated models and methods difficult. In this kind of situation, a simple distributed lags model can be applied. Estimation can be carried out using the simplified Koyck transformation according to which revenue (as a dependent variable) is presented as a function of expenditure and lagged revenue as independent variables (Koyck, 1954). The coefficients of the independent variables can be used to derive IRR. The Koyck transformation makes the estimation look straightforward but actually it is a very challenging task (Franses & van Oest, 2004; Hall, 2007). The estimates of the coefficients are typically very slippery. The objective of the present study is to use this kind of approach to assessing the performance of different estimation methods to derive an estimate of IRR.

There are only few studies considering estimation of IRR for startups. Laitinen (2017) used this kind of model to find different clusters of Finnish startups. He estimated long-term profitability as a form of IRR using the restricted OLS model, where the regression was forced through the origin and an additional restriction on the revenue-expenditure ratio was incorporated in the estimation. Laitinen did not test the performance of the method but reported that for 35% of the startups, estimation was not technically successful. He also concluded that the sensitivity of the estimates made estimation very difficult. Later, Laitinen & Laitinen (2022) presented several methods to estimate the revenue (generation) lag and, simultaneously, IRR for mature Finnish firms. The authors did not test the validity of the estimation methods, because it was not possible to know the correct value of IRR. They concluded their analysis by calling for new approaches to assess estimation methods in an experimental design where the correct values of parameters are given. This study is a response to this call to organize an experimental design to test estimation methods. It is the first study that tests the performance of different estimation methods, when applying the Koyck transformation to derive IRR for startups. The present approach makes it possible to assess the accuracy and dispersion of different estimation methods since the correct IRR is given. Thus, the study provides us with important novel information on how to select the most efficient method for deriving IRR for startups.

The content of this study is comprised of five sections. In the introductory section the motivation and the objective of the study are discussed. The second section presents the framework for the study. First, the generation of the relevant time series (revenue and expenditure) for the experiments is considered. In this context, a similar approach as used by Laitinen (2017) and Laitinen & Laitinen (2022) is applied to describe the non-steady early years of a startup. In the design, different scenarios of time series are classified into nine experiments. Secondly, a framework to estimate IRR using the simplified Koyck transformation is presented. Furthermore, the derivation of the IRR estimate from the coefficients of the independent variables, is discussed. In the third section, the estimation methods are shortly discussed. In all, six different estimation methods are in the nine experiments compared with each other. The fourth section presents the results of the experiments. These results show that the most efficient estimation method is the Ordinary Least Squares (OLS) method where the regression is forced through the origin (RTO). Thus, this method is recommended for deriving IRR for startups also in practice. Finally, the last section summarizes the results of the study.

2. Framework for Experimental Analysis

2.1. Experimental Time-Series

In this study, we will apply a similar business process model as Laitinen & Laitinen (2022) to create the experimental data for the analyses. The objective is, however, to concentrate on simulation of the development of a startup in the first years as did Laitinen (2017). Boumans & Morgan (2001) define this kind of simulation methodology as a hybrid method involving an experiment with an econometric model designed to mimic some process or behavior. This methodology is well-known but rarely applied in startup failure analysis. In testing of the estimation methods, it is important that we know the correct (actual) values of the parameters to assess the performance of alternative methods. Therefore, we need a model that includes the long-term profitability as an input parameter. Following Laitinen (2017) and Laitinen & Laitinen (2022), the business model of this study, firstly, assumes that the expenditures of the startup grow at a constant rate as follows:

$$M_t = M_0 (1+g)^t \tag{1}$$

where M_0 is the initial expenditure in period 0, M_t is expenditure in period t, and g is the constant rate of growth. This assumption simplifies the framework considerably making the analysis mathematically tractable. In spite of this constant growth rate assumption for expenditure, the resulted time series of revenue will in the early stage of a startup be non-steady as it is in practice. In the passage of time, in a more mature stage of the startup, the time series of revenue is expected to approach a steady growth.

Secondly, the business model assumes that there is a fixed cause and effect (causal) relationship between expenditure and generated series of revenue. It is assumed that each periodic expenditure generates an infinite series of revenue with an identical lag structure and the internal rate of return (IRR). It is assumed that the generated series of revenue follow an infinite geometric distribution with a constant lag parameter q. These are quite ordinary assumptions in this kind of distributed lag model of IRR (Luckett, 1984; Feenstra & Wang, 2000; Brief, 2013). In practice, the tail of the revenue distribution converges quickly so that the simplification provides a good approximation. Thus, the non-steady growth of revenues from period 0 to n can be expressed as follows:

$$\begin{aligned} R_n &= KM_0 \sum_{t=0}^{n} (1+g)^t q^{n-t} = KM_n q^n \sum_{t=0}^{n} (1+g)^t q^{-t} \\ &= KM_n \left[\frac{(1+g)^{n+1} - q^{n+1}}{(1+g)^n (1+g-q)} \right] \end{aligned} \tag{2}$$

where K is the level parameter of the lagged revenue distribution (Laitinen, 2017; Laitinen & Laitinen, 2022). Since the series of revenues follows an infinite geometric distribution, the average lag between expenditure and generated revenue is $q/(1-q)$. In the early years, expenditure (1) typically exceeds revenue (2) creating the so-called valley of death for startups. In this period, financial ratios are usually very low.

Equation (2) does not explicitly include IRR to make it possible to estimate long-term profitability. However, the level parameter K gives the revenue contribution realized in the same period as the expenditure is made, as a proportion of this expenditure. Because K is assumed constant, it makes it possible to incorporate IRR or r into the framework. This level parameter K can be solved as a function of q and r in the following way assuming that n will approach infinity:

$$M_n = M_n K \sum_{t=0}^{n} q^t (1+r)^{-t} \to K = \frac{1+r-q}{1+r}, \quad n \to \infty \tag{3}$$

Equation (3) shows that the level K is the higher, the higher is r and the lower is q (the faster revenues are generated by expenditures). In constructing the experimental data of the study, we will use (2) and (3) to control the resulted time series M_t and R_t with respect to g, q, and r. We will construct both these time series for the years 1 - 10 in the early stages of the startup. The challenge of the study is whether we can estimate the long-term profitability *IRR* from these kinds of early non-steady and typically strongly correlating time-series.

Laitinen (2017) has estimated the parameters of a similar distributed lag model for a large set of Finnish startups and used the estimates to cluster startups. For the whole sample, the average estimates in his sample were 0.0392 for the growth of expenditures g, 0.3862 for the lag parameter q, and 0.0445 for the *IRR*, r. On the basis of these values, we selected parameter values 0.04 (g), 0.40 (q), and 0.05 (r) to depict or simulate the time-series in the basic case (experiment). In bankrupt firms, the average parameter values were 0.0055 (g), 0.3468 (q), and −0.0093 (r). In this study, we, however, do not consider bankrupt firms. **Table 1** and **Figure 1** show the time-series for the parameter values used in the basic case. Typically for a startup, revenues grow very slowly in the first years (valley of death) reflecting the non-steady development. However, revenues gradually approach the steady state in the last years when the business process of the startup is getting up in full action. These time series are deterministic although it is clear that actual time-series for startups are stochastic. Typically, the development of startups is non-linear and prone to interruptions and setbacks, which are stochastic and quite difficult to explain using different variables and processes (Garnsey, Stam, & Hefferman, 2006). Therefore, we add a random generator into the model to make the time series of M_t and R_t stochastic. However, we expect that the intrinsic time series when the stochasticity is neglected, behave as being produced by a similar business process as used here to construct the data.

Table 1. Time series data using the parameter values of the basic case (without RAND).

Period	*R*	*M*	*Lag*(*R*)
1	89.14	104.00	61.90
2	102.61	108.16	89.14
3	110.68	112.49	102.61
4	116.69	116.99	110.68
5	121.99	121.67	116.69
6	127.13	126.53	121.99
7	132.31	131.59	127.13
8	137.65	136.86	132.31
9	143.17	142.33	137.65
10	148.90	148.02	143.17

Legend: R = Revenue. M = Expenditure. $Lag(R)$ = lagged revenue. Source: Time series are calculated for the parameter values $r = 0.05$ (IRR), $g = 0.04$ (growth of expenditure), and $q = 0.40$ (revenue lag).

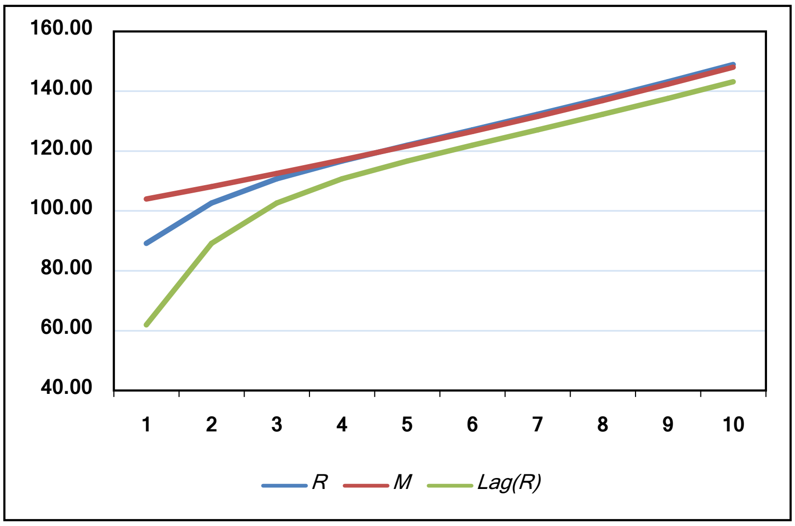

Figure 1. Revenue *R*, expenditure *M*, and lagged revenue *lag*(*R*) in the basic case (without RAND).

The random generator used in the study generates random variables which are used proportionally and applied on each of the yearly figures in the time series. Thus, each value of M_t and R_t ($t = 1$ to 10) will be multiplied by d that is a random variable between 1-RAND and 1+RAND conforming to the uniform distribution. The symmetricity of d makes the means of the time-series equal the original non-stochastic figures reflecting the correct value of r. It is clear that the simulated time series should have similarities with the actual time series. Figure 2 presents the weighted average of the time series for 4207 Finnish startups (source: extracted from Orbis Privat Companies Database, https://www.bvdinfo.com/). The time series in Figure 1 and Figure 2 show strong similarities in spite of the unsteady development of the group of actual startups. In actual startups, the growth of revenues is also non-steady in the first years 1 - 4 but after that valley, revenue grows more steadily together with expenditures. Although there are differences between the modelled and average actual time series, there are also strong similarities reflecting similarity of the business processes. Thus, the experiments can provide us with a realistic insight of the performance of the estimation methods in actual startup data.

The design of this study is based on nine different experiments to test the effects of the control variables g, q, and r. In the basic case, RAND is set equal to 10/1000 which means that d is uniformly distributed between the range from 0.99 to 1.01. Each of the ten figures of the time series M_t and R_t is multiplied by d so that the total effect on the time series is not negligible taking account of that, in the basic case, the periodic growth rate of M_t is only 4%. In the experiments, the parameter values of the model are varied one at the same time so that the experiments describe cases higher/lower profitability, higher/lower growth, higher/lower lag parameter, and higher/lower risk (based on the value of RAND). The parameter values used in the nine experiments are presented in Table 2.

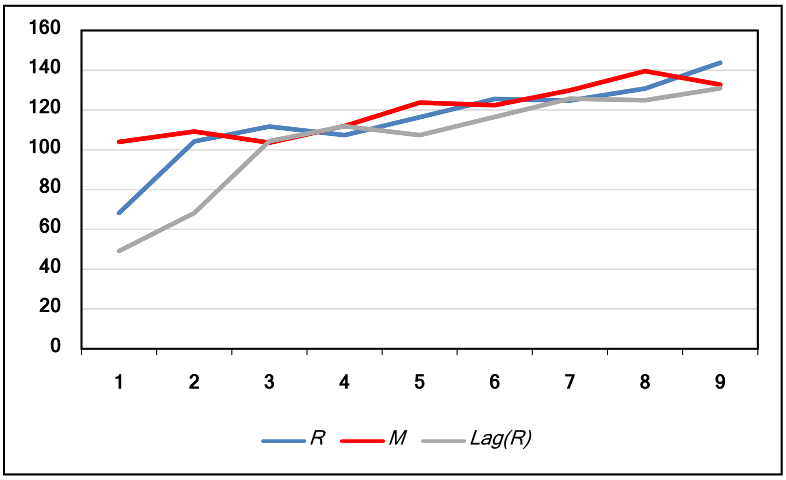

Figure 2. Time series of revenue *R*, expenditure *M*, and lagged revenue *lag*(*R*) as weighted average from 4207 Finnish startups.

Table 2. Parameter values for the nine experimental cases (ceteris paribus analysis).

	Parameter values			
Factor	r	g	q	RAND
1. Basic case	0.05	0.04	0.40	10/1000
2. Higher profitability	0.10	0.04	0.40	10/1000
3. Lower profitability	0.02	0.04	0.40	10/1000
4. Higher growth	0.05	0.08	0.40	10/1000
5. Lower growth	0.05	0.02	0.40	10/1000
6. Higher lag parameter	0.05	0.04	0.60	10/1000
7. Lower lag parameter	0.05	0.04	0.20	10/1000
8. Higher risk	0.05	0.04	0.40	20/1000
9. Lower risk	0.05	0.04	0.40	5/1000

Legend: r = IRR, g = growth rate of expenditure, q = revenue lag parameter, and RAND = random variable. Source: Parameter values in the basic case are based on the average values got by Laitinen (2017). Parameter changes in other cases are based on the selection made by the authors.

The variation of the values is intended to reflect the normal range of the values so that extreme cases are not considered here. The design of the testing is constructed to follow the ceteris paribus (other things being equal) assumption, with respect to both the experiments and the methods (Boumans & Morgan, 2001). This means that the random variables are kept equal, firstly, for each experiment and, secondly, for each method. Thus, for every estimation method and experiment the constructed data are identical to allow a ceteris paribus comparison. It is important to note that we do not present any statistical tests for the estimates but only report how they are behaving in the experiments in the light of descriptive statistics. Especially, we are interested how the derived estimate of *IRR* is behaving when different experiments are considered. This estimate will be derived using the estimates of K and q.

2.2. Estimation of Profitability

The non-steady first years make the estimation task difficult. However, after these few years the independent variables are strongly correlated and bring difficulties in terms of multicollinearity for the estimation of *IRR*. If we assume that the term q^{n+1} in (2) converges quickly after the foundation of the startup, we can approximate the relationship between M_t and R_t using the Koyck (1954) transformation. This transformation makes it possible to estimate the parameters K and q from the simple relationship between R_t, M_t, and lagged R_t (R_{t-1}). Koyck transformation is a procedure used generally in time series analysis to transform an infinite geometric lag model into a model with lagged dependent variable (Franses & van Oest, 2004). In this case, the following simplified equation (without any random terms) is obtained through the transformation:

$$\begin{aligned} & R_t = M_t \cdot K \cdot \frac{1+g}{1+g-q} \rightarrow R_t \cdot \frac{1+g-q}{1+g} = K \cdot M_t \\ & \rightarrow R_t = K \cdot M_t + q \cdot R_{t-1} \end{aligned} \tag{4}$$

Equation (4) shows that we can estimate K and q (and thus r) simultaneously from the time series of R_t and M_t. This linear relationship seems to be intuitively appealing for estimation but the estimation may suffer from serious statistical problems. Firstly, the transformed model is likely to have serial correlation in errors. Secondly, M_t and R_{t-1} are usually strongly correlated leading to obvious difficulties with multicollinearity. The distributed lag model generally gives statistically highly significant estimates. However, these estimates can be remarkably biased (Laitinen & Laitinen, 2022). In this study, the focus is set on how the derived estimates of r are behaving when different statistical methods are applied to the Koyck transformation (4) using artificial data from the early years of startups.

Equation (3) for K shows that using the estimates of K and q, the long-term profitability r can be derived in the following way:

$$r = \frac{1-K-q}{K-1} \tag{5}$$

Thus, r or *IRR* will not be estimated directly but derived indirectly from the estimates of K and q. Factually, this means that an estimation method may give quite accurate estimates of K and q, but still a biased derived estimate for r. This happens if the interplay of the estimates of K and q in terms of Equation (5) is not consistent. If K and q both get too low estimates, $1 - K - q$ can be positive leading to a negative estimate of r although the correct value could be positive.

The estimate of r is in fact quite sensitive to the deviations of K and q from their correct (actual) values. The sensitivity can be assessed using the partial derivatives of r with respect to q and K as follows:

$$\frac{\partial r}{\partial q} = \frac{1}{1-K} \tag{6}$$

$$\frac{\partial r}{\partial K}=\frac{q}{\left(K-1\right)^{2}}$$

The partial derivatives in (6) are both positive which means that if the deviations of K and q from the correct values are positive, the deviation of the derived r from its correct value is more positive. If the deviations of K and q have different signs, they at least partly cancel each other when deriving r.

The total change in r can be assessed by its total derivative, that approximates this change with respect to q and K. In this case, the total derivative of r is the following:

$$\mathrm{d}r=\partial q\frac{1}{1-K}+\partial K\frac{q}{\left(K-1\right)^{2}}\begin{cases}=0 \text{ if } \dfrac{\partial q}{\partial K}=\dfrac{q}{K-1}\\ >0 \text{ if } \dfrac{\partial q}{\partial K}>\dfrac{q}{K-1}\end{cases} \quad (7)$$

This total derivative of r clearly shows that, in order to cancel each other, the deviations of q and K from their correct values must have different signs, since $K < 1$. For the normal values of q and K, their potential effects of deviation on r are relatively similar. For example, if $q = 0.4$ and $K = 0.6$, then the effects of deviations on r cancel each other if they are equal but have different signs. It is important that the estimates of K and q together provide a proper estimate for profitability r. Therefore, we are interested the estimates of K and q separately, but the main interest will be laid on the derived estimate of r.

3. Estimation Methods

In practice, the early time series of startups typically consist a short of non-steady stage (years 1 - 5) and, after the valley of death, a more stable period for survivors towards steady growth path (years 6 - 10). The years in the valley of death are usually financially very poor and reflections of non-steady behavior of revenue time series. In later stages, the behavior of the time series of revenue is usually steadier making, however, the time series strongly correlate over time. Thus, for estimation of long-term profitability, we have typically about ten relevant observations, as in this experimental study, which are partly from a period of non-steady behavior and partly from a period of strongly correlating time-series. Therefore, the nature and length of the data create a challenging environment for estimation. We have only ten observations for estimation, and therefore only simple (regression) estimation methods are applied. For the estimation of the parameters K, q, and r, the following statistical methods are selected:

1) Ordinary Least Squares (OLS) (benchmark);
2) OLS (through the origin, RTO);
3) OLS (restricted);
4) Least Absolute Deviation (LAD) (RTO);
5) Ridge Regression (standardized);
6) Maximum likelihood (ML) (restricted).

3.1. Ordinary Least Squares (OLS)

First, we estimate the Koyck model using the ordinary least squares (OLS). In OLS, the coefficients of the model are estimated by the principle of least squares which means minimizing the sum of the squared differences between the dependent variable and the prediction given by the linear function of the independent variables, in the following way:

$$MSE = \min \sum_{i=1}^{n} \left(y_i - \beta_0 - \beta_1 x_{i1} - \beta_2 x_{i2} \right)^2 \quad (8)$$

where y is R (revenue), x_1 is M (expenditure), x_2 is lagged R (revenue), and n is the number of observations (here, $n = 10$). For assessing the performance of the six estimation methods, OLS will act as a benchmark. In this study, the regression coefficients are solved using the Real Statistics Excel function RegCoeff (X, Y, TRUE) (https://www.real-statistics.com/multiple-regression/multiple-regression-without-intercept/ regression-wo-constant-in-excel/).

3.2. OLS (Through the Origin, RTO)

The simplified form of the Koyck transformation (4) used here does not include any constant (intercept) for estimation. Thus, secondly, we also apply a multiple regression version (OLS) without a constant term (intercept) which is called a regression through the origin (RTO), to minimize the sum of squared errors:

$$\text{MSE} = \min \sum_{i=1}^{n} \left(y_i - \beta_1 x_{i1} - \beta_2 x_{i2} \right)^2 \quad (9)$$

which is comparable with (8) neglecting β_0. Most of the traditional properties for ordinary multiple regression OLS also hold for the RTO regression without the intercept term. Eisenhauer (2003) considers the circumstances where RTO is appropriate or even necessary. RTO may be unavoidable if transformations of the OLS model are needed to correct violations of the Gauss-Markov assumptions. There are also often a priori reasons for believing that $y = 0$ when $x = 0$ and therefore omitting the constant (Eisenhauer, 2003: p. 76). In this study, it is clear that $R = 0$, when $M = 0$ and also lagged $R = 0$. This assumption is also used in constructing the experimental data for testing. The regression coefficients are solved using the Real Statistics Excel function RegCoeff (X, Y, FALSE) where FALSE refers to the regression analysis without the intercept coefficient (RTO) (https://www.real-statistics.com/multiple-regression/multiple-regression-without-intercept/regression-wo-constant-in-excel/).

3.3. OLS (Restricted)

The estimation of the parameters in the Koyck transformation is challenging due to the sensitivity of the estimates but also because they (K and q) must together produce a proper derived estimate of the profitability r. Therefore, it seems to be reasonable to bind the estimates together using some additional restrictions in the estimation. The third method used in this study is therefore the restricted OLS (ROLS), where the restrictions are applied to make the slippery estimates more stable (Laitinen, 2017; Johnston, 1972: pp. 155-159; Fomby, Hill, & Johnson, 1984: pp. 82-85). Equation (4) shows that:

$$\frac{R}{M}=\frac{K(1+g)}{1+g-q}\rightarrow(1+g)K+\frac{R}{M}q=\frac{R}{M}(1+g) \tag{10}$$

which lead to the following matrices of linear restrictions:

$$\boldsymbol{H}=\begin{bmatrix}1 & 0 & 0\\ 0 & 1+g & \frac{R}{M}\end{bmatrix} \text{ and } \boldsymbol{h}=\begin{bmatrix}0\\ \frac{R}{M}(1+g)\end{bmatrix} \tag{11}$$

These restrictions can be presented in a matrix form as $h = H \cdot B$, where B is 3 × 1 matrix of estimates. The first restriction forces the regression through the origin (RTO) whereas the second restriction binds the estimates together using estimates of g and R/M. In this method, g was estimated by OLS from the time series of M_t using the logarithm of M_t as the dependent variable and time index 1 - 10 as the independent variable. The revenue-expenditure ratio R/M was simply calculated summing up separately the last three values of R_t and M_t (observations 8 - 10) and calculating their ratio. The last observations were used, since the time series are all the time becoming steadier to reflect the actual R/M. Using the restrictions in (11) the restricted OLS solutions were calculated using a Visual Basic (VB) program.

3.4. Least Absolute Deviation (LAD) (RTO)

OLS has also characteristics which may weaken its performance. Because the residuals in OLS are squared, MSE gives extensive weight to extreme residuals (quadratic loss function). If the data includes outliers, OLS estimates may be badly distorted. The artificial data of this study has been constructed using the random generator to make uncertainty in observations, which may create unusual errors. In practice, outliers are not rare, since the growth of startups is generally unstable. Therefore, the fourth method used here to estimate the coefficients of the Koyck transformation is the Least Absolute Deviation model (LAD). LAD is based on searching for the minimum sum of absolute errors in regression in the following way:

$$\text{MAE}=\min\sum_{i=1}^{n}\left|y_i-\beta_1 x_{i1}-\beta_2 x_{i2}\right| \tag{12}$$

where MAE is the minimum sum of absolute errors. Thus, LAD is analogous to the least-squares technique of OLS, except that it is based on absolute residuals instead of squared ones. LAD is said to give a more robust set of estimates than OLS, if the data includes outliers. However, LAD may not have a unique solution and an infinite number of different regression lines may lead to an identical MAE. When the errors follow the Laplace distribution, the absolute errors estimates are maximum likelihood and hence asymptotically efficient (Thanoon, 2015). There are different algorithms to find efficient estimates for the coefficients of the model (Chen & Derezinski, 2021). In this study, we use the method of iteratively reweighted least squares (IRLS) to solve the LAD estimates. IRLS is an iterative method in which each step involves solving a weighted least squares problem to find the maximum likelihood estimates of a generalized linear model. The process is here automated using the Real Statistics LAD Regression function LADRegCoeff (*X*, *Y*, FALSE, 100) where FALSE refers to the regression without an intercept and 100 is the number for iterations (25 is the default value) (https://www.real-statistics.com/multiple-regression/lad-regression/lad-regression-analysis-tool/)

3.5. Ridge Regression (Standardized)

The fifth estimation method applied in this study is the Ridge regression (RR) that is a method of estimating the coefficients of multiple-regression models when the independent variables are highly correlated as in the data of this study. Following the Gauss-Markov Theorem OLS gives regression coefficients which are unbiased with the least variance. However, when the independent variables are strongly correlated leading to multicollinearity, these coefficients are slippery even to small changes in *X*. In this kind of situation, the variance of RR model (sum of squared errors, SSE) can be lower than that of OLS estimates but the estimates are biased. In the Ridge regression, bias is artificially incorporated to the regression equation to make the variance of the estimates lower. For RR, SSE can be presented in this analysis as follows:

$$\mathrm{SSE} = \sum_{i=1}^{n}\left(y_i - \beta_1 x_{i1} - \beta_2 x_{i2}\right)^2 + \lambda\sum_{j=1}^{2}\beta_j^2 = \left(Y - XB\right)^{\mathrm{T}}\left(Y - XB\right) + \lambda B^{\mathrm{T}} B \quad (13)$$

where lambda (λ) is a non-negative ridge parameter reflecting the bias.

RR estimates are biased as lambda increases but may give more precise estimates than OLS. If lambda = 0, RR returns to OLS. If lambda increases towards infinity, all the regression coefficients shrink towards zero. The essential question in RR is, how to select lambda to get best estimates for the coefficients. It has been suggested that the value of lambda should be small enough such that the SSE of RR is less than the SSE of OLS (Hoerl, Kennard, & Baldwin, 1975). Researchers have proposed different techniques of estimation for the lambda parameter (Lukman & Ayinde, 2016; Duzan & Shariff, 2015). In this study, we use RidgeRegCoeff (*X*, *Y*, lambda, TRUE) to solve the standardized coefficients (without an intercept) for M_t and $R_{t\text{-}1}$ (https://www.real-statistics.com/multiple-regression/ridge-and-lasso-regression/ridge-regression-example/). The lambda parameter is solved simply by the Real Statistics function RidgeLambda (*X*, 5, 100) that gives the lowest lambda value for RR in *X* that generates a maximum VIF value less than 5 (default VIF value). The function is allowed to use 100 iterations in the search (default value is 25).

3.6. Maximum Likelihood (ML) (Restricted)

The parameters of a Koyck linear regression model can be estimated also using a maximum likelihood estimation (MLE) procedure. MLE is a probabilistic framework for finding the probability distribution and parameter estimates that best describe the observed data. MLE differs from OLS in that OLS does not make any assumption of the probabilistic nature of the variables and is, therefore, regarded as deterministic whereas MLE is probabilistic by nature. In MLE, the parameters are estimated by maximizing a likelihood function in the way that the observed data is most probable. The estimate that maximizes the likelihood function is called the ML estimate. In this study, we estimate the parameters of the Koyck model that in the general form can be presented as follows (Franses & van Oest, 2004):

$$R_t = \beta_0 + \beta_1 K_t + \beta_2 R_{t-1} + \varepsilon_t - \beta_2 \varepsilon_{t-1} \tag{14}$$

which is called an ARMAX model. The autoregressive (AR) part concerns R_{t-1}, the moving average part (MA) concerns ε_{t-1} and the explanatory variables part (X) concerns K_t. In this form, the Koyck model is complicated and iterative methods are needed to solve the ML estimates.

However, if we assume that ε_t is normally distributed with a zero mean and variance σ^2, the ML estimates can be found maximizing the following conditional likelihood function (Klein, 1957):

$$L = -n\left(\log\sqrt{2\pi} + \log\sigma\right) - \frac{1}{2\sigma^2}\sum_{t=1}^{n}\varepsilon_t^2 \tag{15}$$

where n is the number of observations. In this simplification, to maximize the likelihood function L is equivalent to minimize the squared sum of errors which means that the ML estimates for β_1 (K) and β_2 (q) are equivalent to the OLS estimates. These assumptions mean that we have neglected the MA part of (14) making it possible to estimate the parameters by OLS. Franses & van Oest (2004) criticize this reduced form of the model, since ε_{t-1} and R_{t-1} may not be uncorrelated which can lead to a downward bias of β_2 (q). If we add linear restrictions to the model, the OLS and ML estimates are not equivalent anymore, since they differ with respect to the restriction adjustment. The likelihood function of the restricted model is differentiable and the first-order conditions can be solved analytically to find the ML estimates. In this study, we calculated the ML estimates using a Visual Basic (VB) program for the reduced Koyck model where the restrictions (11) were included. It is expected that the estimates are correlated with the OLS estimates, since they differ only with respect to the adjustment.

4. Estimation Results

There are many interesting points in the evaluation of the results from the experiments. Firstly, we can compare the performance of the unstandardized OLS (OLSU) and the Ridge Regression (RR) to reveal the effect of adding the bias (lambda). If lambda = 0, these methods give the same result. Secondly, we can assess the effect of forcing the regression through the origin (neglecting the intercept) by comparing the performance of OLSU and RTO OLS (OLSRT). Thirdly, the effect of the revenue-expenditure restriction can be found when the results from OLSRT and the restricted OLS (ROLS) are compared with each other. Fourthly, the impact of the minimization of the sum of absolute errors instead of the squared ones can be showed by comparing the performance of OLSRT and RTO LAD (LADRT). Sixthly, the comparison of the results from ROLS and the restricted MLE (RMLE) gives an insight how efficient MLE is, when compared with OLS. In this modelling, only the restriction adjustments make the differences between these estimation methods. In general, the purpose is of course to rank the methods with respect to their overall estimation performance divided as 1) accuracy and 2) dispersion.

4.1. Basic Case

In the basic case, the average parameters from Finnish startups are used so that $r = 0.05$, $g = 0.04$, and $q = 0.40$. These values imply that the level parameter $K = 0.6190$. In this case, uncertainty is not set at a very high level, since RAND = 10. For the 100 rounds run using the random generator, the average correlation between the independent variables M_t and R_{t-1} is 0.935 leading to an average VIF = 7.973 for both variables. Thus, there is multicollinearity in the estimation although VIF is less than 10 that is sometimes regarded as the acceptable level. When the restriction VIF = 5 (usual standard) is set, the Real Statistics function returns 0.157 as the lowest average lambda value (for Ridge Regression) that generates a maximum VIF value less than 5. This lambda value is used in RR to reflect bias. In general, multicollinearity is not very high which may make the effect of the bias on the estimates relatively small. The descriptive statistics of the estimation results from the experiment are presented in Table 3.

Table 3. Descriptive statistics of the estimation results from the basic case.

	OLSU	OLSRT	ROLS	LADRT	RR	RMLE
Panel 1. Estimates of K (0.6190)						
Mean	0.6192	0.6218	0.6208	0.6210	0.6190	0.6203
Standard Deviation	0.0521	0.0160	0.0165	0.0196	0.0520	0.0404
Variation Coefficient	0.0841	0.0257	0.0265	0.0315	0.0840	0.0651
10% percentile	0.5565	0.6014	0.6007	0.5979	0.5563	0.5756
25% percentile	0.5880	0.6099	0.6093	0.6096	0.5878	0.5945
Median	0.6208	0.6220	0.6193	0.6207	0.6206	0.6200
75% percentile	0.6567	0.6331	0.6322	0.6348	0.6564	0.6486
90% percentile	0.6779	0.6413	0.6405	0.6484	0.6775	0.6691
Minimum value	0.4703	0.5704	0.5780	0.5665	0.4702	0.5052
Maximum value	0.7328	0.6617	0.6675	0.6657	0.7323	0.7101

Continued

Mean error	0.0002	0.0028	−0.0018	−0.0018	−0.0019	−0.0011
Mean absolute error	0.0409	0.0130	0.0150	0.0173	0.0237	0.0312
Number of estimates > correct value	51	56	48	48	45	47
Number of estimates < correct value	49	44	52	52	55	53
Panel 2. Estimates of *q* (0.40)						
Mean	0.3979	0.3970	0.3982	0.3982	0.3981	0.3989
Standard Deviation	0.0298	0.0175	0.0187	0.0219	0.0298	0.0402
Variation Coefficient	0.0749	0.0441	0.0470	0.0551	0.0747	0.1008
10% percentile	0.3657	0.3759	0.3780	0.3707	0.3659	0.3523
25% percentile	0.3773	0.3856	0.3871	0.3836	0.3775	0.3706
Median	0.3960	0.3965	0.3984	0.3992	0.3961	0.3984
75% percentile	0.4183	0.4104	0.4116	0.4124	0.4185	0.4238
90% percentile	0.4334	0.4189	0.4199	0.4243	0.4335	0.4470
Minimum value	0.3303	0.3525	0.3436	0.3479	0.3305	0.3096
Maximum value	0.4828	0.4547	0.4438	0.4606	0.4828	0.5163
Mean error	−0.0021	−0.0030	−0.0018	−0.0018	−0.0019	−0.0011
Mean absolute error	0.0237	0.0143	0.0150	0.0173	0.0237	0.0312
Number of estimates > correct value	45	42	48	48	45	47
Number of estimates < correct value	55	58	52	52	55	53
Panel 3. Estimates (derived) of *r* (0.05)						
Mean	0.0546	0.0496	0.0499	0.0504	0.0542	0.0511
Standard Deviation	0.0764	0.0053	0.0083	0.0080	0.0762	0.0097
Variation Coefficient	1.3989	0.1066	0.1654	0.1589	1.4051	0.1902
10% percentile	−0.0272	0.0430	0.0400	0.0417	−0.0274	0.0400
25% percentile	−0.0043	0.0458	0.0445	0.0463	−0.0046	0.0440
Median	0.0428	0.0498	0.0502	0.0501	0.0424	0.0509
75% percentile	0.1009	0.0532	0.0561	0.0555	0.1003	0.0566
90% percentile	0.1572	0.0564	0.0598	0.0596	0.1566	0.0627
Minimum value	−0.0989	0.0308	0.0213	0.0174	−0.0990	0.0253
Maximum value	0.2613	0.0633	0.0721	0.0703	0.2599	0.0881
Mean error	0.0046	−0.0004	−0.0001	0.0004	0.0042	0.0011
Mean absolute error	0.0602	0.0042	0.0064	0.0060	0.0600	0.0075
Number of estimates > correct value	47.0000	48.0000	53.0000	51.0000	46.0000	54.0000
Number of estimates < correct value	53.0000	52.0000	47.0000	49.0000	54.0000	46.0000

Legend: Correct values of the parameters are presented in parentheses. OLSU = OLS (unstandardized coefficients); OLSRT = OLS forced through the origin (RTO); ROLS = OLS (RTO) with revenue-expenditure restriction; LADRT = LAD forced through the origin (RTO); RR = Ridge Regression (standardized coefficients); RMLE = ML (RTO) with revenue-expenditure restriction. Source: Descriptive statistics calculated by the authors.

Panel 1 presents the descriptive statistics for the estimates of K (correct value is 0.6190). On average, all estimation methods give a quite accurate results so that the median of the estimates slightly exceeds the correct value. However, there are significant differences in the dispersion of the estimates. The estimates given by OLSU, RR, and RMLE clearly show a higher variation than estimates got by LADRT, ROLS, and OLSRT. The variation is largest for OLSU so that the minimum value is 0.470 and the maximum value is 0.733 reflecting a wide range of the estimate. The estimates of RR are very close to those of OLSU but anyway the bias has corrected the estimates to the right direction. It seems that the rough order of overall performance in estimation is ROLS, LADRT, OLSRT, RMLE, RR, and OLSU.

Panel 2 presents the results of the estimation of q (the correct value is 0.400). For each estimation method, the estimates are on average close to the correct value but little below it. This means that when the estimates of K on average exceed the correct value, the errors partly cancel each other in (derived) estimation of r. The results can indicate that the neglection of the MA part of the original Koyck model may have led to a downward bias as commented by Franses & van Oest (2004). The highest dispersions of the estimates are found for RMLE, OLSU, and RR while OLSRT, ROLS, and LADRT has the lowest ones. For RMLE the minimum estimate is as low as 0.031 whereas the maximum is 0.516 leading to a large range of errors. The estimates for OLSU and RR are again very close to each other in the way that the bias (lambda) has corrected the estimates to the right direction. The estimation methods can be set in the following order of performance: ROLS, OLSRT, LADRT, RR, OLSU, and RMLE.

Panel 3 shows the results for the derived profitability r (0.05 is the correct value). The means of the estimates are close to the correct value with an exception for OLSU and RR, which give estimates that clearly exceed it. However, the medians of the estimates for these methods greatly underestimate r reflecting an asymmetric distribution of errors. The variations of the estimates for them are exceptionally high so that the mean absolute difference to the correct value is 0.060 for both methods. Thus, in practice, these methods are useless in estimation of profitability. OLSRT, ROLS, and LADRT show a quite equal high performance whereas RMLE shows a lower one. The mean absolute errors of the estimates for OLSRT, ROLS, and LADRT are very small referring to a high performance. The rough order of performance of the estimation methods can be set in the following way: OLSRT, ROLS, LADRT (very good methods), RMLE, OLSU, and RR.

In summary, the results show that the bias in RR does not improve the performance of OLSU. The revenue-expenditure restriction in ROLS does not increase the efficiency of OLSRT that seems to be the most efficient method slightly beating LADRT. RMLE is more efficient in estimating r than in estimating K and q which may be due to the revenue-expenditure restriction that ties the estimates together. **Table 4** shows that the estimates of r given by OLSU and

RR are fully correlated (1.000) due to the weak effect of lambda on RR estimates. In the same way, the estimates given by RMLE and ROLS are strongly correlated (0.968) maybe because these methods use similar restriction adjustments built on the OLS estimate. In the estimation of *r*, the methods can be classified into three groups on the basis of their efficiency: excellent or good methods (OLSRT, ROLS, and LADRT), satisfactory methods (RMLE), and poor methods (OLSU and RR). For the 100 experimental rounds, Figure 3 demonstrates the significant differences in the estimates between the most efficient method OLSRT and OLSU, that is one of the most inefficient methods. The estimates of OLSRT for the rounds are without exception very close to the correct value of *r* (0.05). However, the estimates of OLSU have a very large dispersion around the correct value. The variation of the estimates given by OLSU is so large that it makes this method useless in practice. It also shows the crucial effect of forcing the regression through the origin on the estimates.

Table 4. Pearson correlation coefficients between the profitability (*r*) estimates given by the different estimation methods in the basic case.

	OLSU	OLSRT	ROLS	LADRT	RR	RMLE
OLSU	1.0000	0.2451	0.6837	0.1060	1.0000	0.7944
OLSRT	0.2451	1.0000	0.7217	0.7564	0.2455	0.6309
ROLS	0.6837	0.7217	1.0000	0.5137	0.6839	0.9675
LADRT	0.1060	0.7564	0.5137	1.0000	0.1064	0.4194
RR	1.0000	0.2455	0.6839	0.1064	1.0000	0.7944
RMLE	0.7944	0.6309	0.9675	0.4194	0.7944	1.0000

Legend: For the abbreviations of the estimation methods see Table 3. Source: Pearson correlation coefficients calculated by the authors.

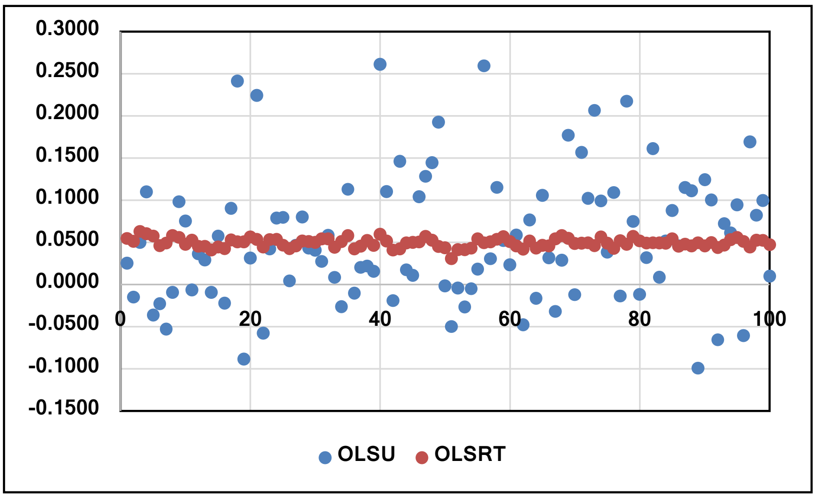

Figure 3. The dispersion of the profitability (*r*) estimates given by OLSU and OLSRT in 100 rounds (the basic case). Legend: For the abbreviations of the estimation methods see Table 3.

4.2. Experimental Cases

4.2.1. Estimate of *K*

Table 5 presents a summary of the descriptive statistics for the estimates of *K* with respect to accuracy and dispersion. Panel 1 presents the mean of the estimates by method and experiment. This panel shows that the methods generally slightly overestimate *K*, if the means of the estimates are considered. OLSRT overestimates *K* in the first seven and LADRT in the first eight experiments reflecting the effect of the restriction (intercept = 0). In the high and low RAND circumstances, almost all methods underestimate *K*. RR underestimates *K* in six experiments, only high *g* and high *q* circumstances lead to a very small overestimation. For the high RAND circumstances, underestimation is significant. In general, for OLSU and RR, the means of the estimates are surprisingly close to the correct value.

Panel 2 of the table presents the variation coefficients of the estimates by method and experiment. In all experiments, the ranks of the methods are identical. RR leads a slightly smaller variation than OLSU, but both methods have clearly the largest variations in each experiment. In the ranking of methods with respect to the variation coefficient, OLSTR shows the lowest variation followed by ROLS, LADTR, and, with a quite high variation, RMLE. The high RAND circumstances significantly increase variation especially for OLSU, RR, and RMLE. For OLSTR, ROLS, and LADTR also high *q* circumstances tends to increase variation.

Panel 3 presents the mean absolute errors of the estimates. For this central measure of accuracy, the ranking of the methods is clear-cut. In each experiment, OLSTR leads to the smallest error closely followed by ROLS and LADTR. Clearly, the largest errors are given by OLSU whereas the second largest errors are got by RMLE. For this measure of accuracy, RR tends to significantly improve the estimates when compared with OLSU. Thus, the incorporation of bias (lambda) in the model seems to shrink the absolute error. It is notable that the mean absolute errors for each method are quite same in the three first experiments (basic case, high *r*, and low *r*). Thus, the height of *r* does not affect the absolute error in the estimate of *K*. However, for the low *q* circumstances the errors are very large.

4.2.2. Estimate of *q*

Table 6 presents the descriptive statistics of the estimates of *q* with respect to experiment and estimation method. Panel 1 of the table presents the means of the estimates showing that all methods slightly underestimate *q* in all of the six first experiments. The only exception is found for RMLE in the low *g* circumstances. This underestimation may be due to the neglection of the MA part of the Koyck model (Franses & van Oest, 2004). In the low *q* circumstances, OLSU, ROLS, RR and RMLE lead to overestimation while in the high RAND circumstances it happens to OLSU, OLSRT, RR, and RMLE. However, in the low RAND circumstances all methods lead to overestimation of *q*. All six methods give the same results for the first three experiments (basic case, high *r*, and low *r*) indicating that the height of profitability does not affect estimation. For these experiments, RMLE gives the average estimate closest to the correct value, followed by ROLS, LADTR, and RR. In the low RAND circumstances, every method leads to overestimation. Generally, the differences in the average estimates between the methods are quite small.

Table 5. Descriptive statistics on the estimates of K (the level parameter of revenue distribution).

Experiment	OLSU	OLSRT	ROLS	LADRT	RR	RMLE	Correct value of K
Panel 1. Mean of estimate							
Basic case	0.6192	0.6218	0.6208	0.6210	0.6190	0.6203	0.6190
High r (0.10)	0.6365	0.6392	0.6381	0.6383	0.6363	0.6376	0.6364
Low r (0.02)	0.6080	0.6106	0.6095	0.6097	0.6078	0.6090	0.6078
High g (0.08)	0.6210	0.6217	0.6204	0.6210	0.6209	0.6214	0.6190
Low g (0.02)	0.6119	0.6219	0.6209	0.6209	0.6119	0.6159	0.6190
High q (0.60)	0.4295	0.4302	0.4331	0.4303	0.4293	0.4290	0.4286
Low q (0.20)	0.8023	0.8101	0.8053	0.8130	0.8012	0.8046	0.8095
High RAND (20)	0.5940	0.6187	0.6188	0.6228	0.5938	0.6013	0.6190
Low RAND (5)	0.6138	0.6182	0.6186	0.6176	0.6135	0.6150	0.6190
Panel 2. Variation coefficient							
Basic case	0.0841	0.0257	0.0265	0.0315	0.0840	0.0651	
High r (0.10)	0.0841	0.0257	0.0265	0.0315	0.0840	0.0656	
Low r (0.02)	0.0841	0.0257	0.0265	0.0315	0.0840	0.0647	
High g (0.08)	0.0698	0.0266	0.0278	0.0316	0.0697	0.0605	
Low g (0.02)	0.1225	0.0253	0.0259	0.0312	0.1224	0.0828	
High q (0.60)	0.1138	0.0159	0.0154	0.0236	0.1137	0.0800	
Low q (0.20)	0.0994	0.0508	0.0533	0.0565	0.0991	0.0862	
High RAND (20)	0.1753	0.0506	0.0522	0.0566	0.1752	0.1323	
Low RAND (5)	0.0362	0.0135	0.0144	0.0162	0.0362	0.0280	
Panel 3. Mean absolute error							
Basic case	0.0409	0.0130	0.0150	0.0173	0.0237	0.0312	
High r (0.10)	0.0420	0.0133	0.0150	0.0173	0.0237	0.0314	
Low r (0.02)	0.0401	0.0127	0.0150	0.0173	0.0237	0.0310	
High g (0.08)	0.0344	0.0132	0.0165	0.0184	0.0279	0.0315	
Low g (0.02)	0.0595	0.0128	0.0142	0.0166	0.0218	0.0386	
High q (0.60)	0.0390	0.0055	0.0083	0.0101	0.0169	0.0275	
Low q (0.20)	0.0617	0.0333	0.0386	0.0389	0.0474	0.0537	
High RAND (20)	0.0884	0.0258	0.0297	0.0304	0.0488	0.0671	
Low RAND (5)	0.0182	0.0067	0.0080	0.0095	0.0106	0.0142	

Legend: For the abbreviations of the estimation methods see Table 3. r = IRR, g = growth rate of expenditure, q = revenue lag parameter, and RAND = random variable. Source: From experimental values summarized by the authors.

Table 6. Descriptive statistics on the estimates of *q*.

Experiment	OLSU	OLSRT	ROLS	LADRT	RR	RMLE	Correct value of *q*
Panel 1. Mean of estimate							
Basic case	0.3979	0.3970	0.3982	0.3982	0.3981	0.3989	0.4000
High *r* (0.10)	0.3979	0.3970	0.3982	0.3982	0.3981	0.3989	0.4000
Low *r* (0.02)	0.3979	0.3970	0.3982	0.3982	0.3981	0.3989	0.4000
High *g* (0.08)	0.3972	0.3970	0.3985	0.3981	0.3974	0.3976	0.4000
Low *g* (0.02)	0.3995	0.3970	0.3981	0.3983	0.3995	0.4034	0.4000
High *q* (0.60)	0.5983	0.5981	0.5935	0.5981	0.5984	0.5978	0.6000
Low *q* (0.20)	0.2032	0.1994	0.2044	0.1969	0.2040	0.2054	0.2000
High RAND (20)	0.4111	0.4002	0.3993	0.3956	0.4112	0.4180	0.4000
Low RAND (5)	0.4031	0.4010	0.4005	0.4018	0.4032	0.4042	0.4000
Panel 2. Variation coefficient							
Basic case	0.0749	0.0441	0.0470	0.0551	0.0747	0.1008	
High *r* (0.10)	0.0749	0.0441	0.0470	0.0551	0.0747	0.1016	
Low *r* (0.02)	0.0749	0.0441	0.0470	0.0551	0.0747	0.1003	
High *g* (0.08)	0.0877	0.0484	0.0522	0.0581	0.0876	0.1000	
Low *g* (0.02)	0.0688	0.0422	0.0446	0.0529	0.0687	0.1227	
High *q* (0.60)	0.0349	0.0135	0.0137	0.0206	0.0349	0.0577	
Low *q* (0.20)	0.3014	0.2197	0.2278	0.2501	0.2990	0.3417	
High RAND (20)	0.1398	0.0854	0.0915	0.0972	0.1396	0.1900	
Low RAND (5)	0.0323	0.0232	0.0259	0.0284	0.0322	0.0426	
Panel 3. Mean absolute error							
Basic case	0.0237	0.0143	0.0150	0.0173	0.0237	0.0312	
High *r* (0.10)	0.0237	0.0143	0.0150	0.0173	0.0237	0.0314	
Low *r* (0.02)	0.0237	0.0143	0.0150	0.0173	0.0237	0.0310	
High *g* (0.08)	0.0279	0.0154	0.0165	0.0184	0.0279	0.0315	
Low *g* (0.02)	0.0218	0.0137	0.0142	0.0166	0.0218	0.0386	
High *q* (0.60)	0.0169	0.0067	0.0083	0.0101	0.0169	0.0275	
Low *q* (0.20)	0.0476	0.0354	0.0386	0.0389	0.0474	0.0537	
High RAND (20)	0.0489	0.0283	0.0297	0.0304	0.0488	0.0671	
Low RAND (5)	0.0106	0.0074	0.0080	0.0095	0.0106	0.0142	

Legend: For the abbreviations of the estimation methods see Table 3. *r* = IRR, *g* = growth rate of expemditure, *q* = revenue lag parameter, and RAND = random variable. Source: From experimental values summarized by the authors.

Panel 2 presents the variation coefficients of the estimates. With respect to this dispersion measure the ranking of the estimation methods is clear. In each experiment, the lowest variation coefficient has got by OLSTR followed by ROLS and LADTR. The highest variation coefficients have produced by RMLE (highest), OLSU, and RR. Generally, the coefficients are logically lowest in the low RAND circumstances and exceptionally high in the high RAND circumstances. However, these variation measures are clearly highest in the low *q* circumstances. Panel 3 presents the mean absolute errors for the estimation methods. Generally, the errors are the smallest for OLSTR, the second smallest for ROLS, the third smallest for LADTR, followed by RR, OLSU, and RMLE. The errors are exceptionally high in the low *q* and the high RAND circumstances. Similarly, they are exceptionally low in the low RAND circumstances. Thus, uncertainty in the form of RAND is an important source of the size of errors.

4.2.3. Estimate of *r*

Table 7 presents descriptive statistics of the (derived) estimates for *r* by estimation method and experiment. For this study, these derived estimates are the most important estimates. Panel 1 presents the means of the (derived) estimates which in general are quite close to the correct value. In most circumstances, the estimates given by OLSU, LADTR, RR, and RMLE overestimate *r* whereas the estimates given by OLSRT and ROLS underestimate it. In the circumstances of high *r*, all methods slightly overestimate profitability. In the circumstances of low *q*, RR and OLSU give extremely low estimates whereas RMLE leads to an estimate clearly overestimating *r*.

Panel 2 of the table presents the variation coefficients of the estimates by estimation method and experiment. With respect to this dispersion measure, the ranking of the estimation methods is clear. The smallest coefficient of variation in each experiment has got by OLSTR followed by LADTR, ROLS, and RMLE. However, OLSU and RR have got exceptionally high coefficients in each experiment, with an exception for the circumstances of the low *q*, where the coefficients are negative (due to a negative mean). With an exception for these negative values, all estimation methods have got a very high variation coefficient in the circumstances of the low *r*, the low *q*, and the high RAND. In practice, OLSU and RR are not useful estimation methods due to very high dispersion.

Panel 3 presents the mean absolute errors in the (derived) estimation of *r* for each estimation method and experiment. The ranking of estimation methods with respect to this measure is clear-cut. The lowest mean of the absolute error has provided by OLSRT whereas the second lowest mean has got by LADRT and the third lowest by ROLS. In this ranking, OLSU and RR are clearly the last ones with very high mean absolute errors. The estimation methods have got the highest absolute errors in the circumstances of the low *q* and the high RAND. RMLE gives satisfactory errors but is clearly beaten by OLSRT, LADRT, and ROLS. However, RMLE has produced quite low errors especially in the circumstances of the low *r* and the high *g*, where *g* exceeds *r* (and steady ratio of revenue to expenditure exceeds unity).

Table 7. Descriptive statistics on the estimates of r (IRR).

Experiment	OLSU	OLSRT	ROLS	LADRT	RR	RMLE	Correct value of r
Panel 1. Mean of estimate							
Basic case	0.0546	0.0496	0.0499	0.0504	0.0542	0.0511	0.0500
High r (0.10)	0.1075	0.1004	0.1005	0.1011	0.1070	0.1026	0.1000
Low r (0.02)	0.0232	0.0192	0.0196	0.0200	0.0228	0.0204	0.0200
High g (0.08)	0.0516	0.0490	0.0496	0.0500	0.0515	0.0501	0.0500
Low g (0.02)	0.0580	0.0499	0.0502	0.0507	0.0579	0.0516	0.0500
High q (0.60)	0.0534	0.0496	0.0468	0.0498	0.0531	0.0470	0.0500
Low q (0.20)	−0.0756	0.0495	0.0479	0.0518	−0.2620	0.0630	0.0500
High RAND (20)	0.0584	0.0493	0.0469	0.0483	0.0579	0.0509	0.0500
Low RAND (5)	0.0454	0.0504	0.0499	0.0508	0.0450	0.0500	0.0500
Panel 2. Variation coefficient							
Basic case	1.3989	0.1066	0.1654	0.1589	1.4051	0.1902	
High r (0.10)	0.8607	0.0557	0.0787	0.0765	0.8618	0.1740	
Low r (0.02)	2.9160	0.3164	0.4616	0.4523	2.9528	0.3542	
High g (0.08)	0.6920	0.1487	0.2095	0.2089	0.6924	0.1579	
Low g (0.02)	2.7673	0.0963	0.1507	0.1364	2.7676	0.2617	
High q (0.60)	1.0571	0.0594	0.0791	0.0962	1.0592	0.0902	
Low q (0.20)	−26.2703	0.2713	0.5007	0.3451	−11.6643	0.8726	
High RAND (20)	3.7435	0.2167	0.3348	0.3363	3.7587	0.4250	
Low RAND (5)	0.7174	0.0575	0.0889	0.0843	0.7211	0.0988	
Panel 3. Mean absolute error							
Basic case	0.0602	0.0042	0.0064	0.0060	0.0600	0.0075	
High r (0.10)	0.0724	0.0043	0.0061	0.0058	0.0722	0.0142	
Low r (0.02)	0.0534	0.0050	0.0071	0.0067	0.0533	0.0057	
High g (0.08)	0.0289	0.0059	0.0082	0.0077	0.0288	0.0062	
Low g (0.02)	0.1207	0.0038	0.0059	0.0051	0.1206	0.0105	
High q (0.60)	0.0449	0.0024	0.0039	0.0038	0.0449	0.0042	
Low q (0.20)	0.4532	0.0104	0.0175	0.0129	0.5848	0.0281	
High RAND (20)	0.1397	0.0086	0.0125	0.0122	0.1393	0.0153	
Low RAND (5)	0.0262	0.0024	0.0035	0.0035	0.0262	0.0039	

Legend: For the abbreviations of the estimation methods see Table 3. r = IRR, g = growth rate of expenditure, q = revenue lag parameter, and RAND = random variable. Source: From experimental values summarized by the authors.

5. Summary

The objective of this study was to compare the performance of alternative estimation methods in estimating the long-term profitability (*r*) of startups. The question is important, since most startups in the early stages suffer from a valley of death where the financial ratios are very low and do not predict future success properly. The distributed lag model based on the simplified Koyck transformation provides us with an approach to assess profitability of startups even from a short time-series. The objective was to use experimental design to assess different estimation methods used to estimate the parameters of the model, especially *r*. This kind of approach is useful because it makes it possible to include the actual profitability in the experimental data and, therefore, assess the measurement error.

The model used to construct the experimental data was based on the causal relationship between expenditures and revenues. It was assumed that a startup yearly invests an amount of expenditures that grow at a steady rate *g*. These yearly expenditures are assumed to create an infinite flow of geometrically distributed revenues with a lag parameter *q* and a constant *r*. Thus, the experiments were controlled by three parameters, *g*, *q*, and *r*. For each parameter two levels were constructed (High/Low) leading to $2^3 = 8$ experiments, in all 9 with the basic case (experiment). In each experiment, randomized data were used to estimate the three parameters of the Koyck model, the level parameter *K*, the lag parameter *q*, and, finally, the derived parameter *r* (calculated from *K* and *q*). Six different estimation methods (OLSU, OLSRT, ROLS, LADRT, RR, and RMLE) were assessed using descriptive statistics from the estimation in different experiments (circumstances).

The estimation methods generally slightly overestimated *K* and slightly underestimated *q*. The underestimation of *q* may be a result of the neglection of the MA part of the Koyck model in estimation (Franses & van Oest, 2004). For *r*, the estimates were generally accurate with a slight overestimation (OLSU, LADTR, RR, and RMLE) or underestimation (OLSTR & ROLS). In almost all experiments, the ranking of the six methods with respect to accuracy and dispersion was clear-cut: OLSTR (best), LADRT, ROLS, RMLE, RR, and OLSU (worst). The most challenging circumstances of the estimation methods were created by the experiment of the low *q*, where *q* was set to 0.2 to describe the circumstances of a labor-intensive startup or a retail startup with a very quick generation of revenues. When $q = 0.2$, the average time lag between expenditure and revenue is only 0.25 year making the time-series of revenue and expenditure almost totally overlapping. **Figure 4** shows the time-series (without RAND) for this experiment reflecting a very difficult estimation task. The time series of revenue is non-steady only in years 1 - 2 leading to a very short valley of death, and after that there is very little information between revenue, expenditure, and lagged revenue to be used in the estimation of IRR.

Table 8 presents a summary of the assessment of the six estimation methods for the estimation of the long-term profitability *r*. For each of the nine experiments, the table shows the rough ranks of the methods with respect to the efficiency of estimation. The ranks indicate that OLSRT is the superior estimation method. In the experiments of the low r and high g (where $r < g$), OLSRT underestimated r and was therefore slightly beaten by RMLE. In these experiments only, the (steady) revenue-expenditure ratio was less than unity, because $r < g$. However, also in these experiments the differences in efficiency between the top methods were very small and OLSRT anyway gave the smallest absolute error in estimation. However, when analyzing startups with low profitability or high growth rate, it would be safe to apply also RMLE to check whether there is found underestimation of r. It should be noted that also ROLS, that has a revenue-expenditure restriction like RMLE, gave efficient estimates without any underestimation. Thus, the restriction adjustment seems to be useful especially in these kinds of startups (circumstances) where $g > r$.

Table 8. The ranks of different estimation methods by experiment.

Experiment	OLSU	OLSRT	ROLS	LADRT	RR	RMLE
Basic case	6	1	2	2	5	4
High r (0.10)	5	1	2	2	6	4
Low r (0.02)	6	2	3	3	5	1
High g (0.08)	5	2	3	3	5	1
Low g (0.02)	5	1	2	2	5	4
High q (0.60)	5	1	2	3	6	4
Low q (0.20)	5	1	3	2	6	4
High RAND (20)	5	1	3	2	5	4
Low RAND (5)	5	1	2	3	6	4
Sum of ranks	47	11	22	22	49	30

Legend: For the abbreviations of the estimation methods see Table 3. r = IRR, g = growth rate of expenditure, q = revenue lag parameter, and RAND = random variable. Source: From experimental values summarized by the authors.

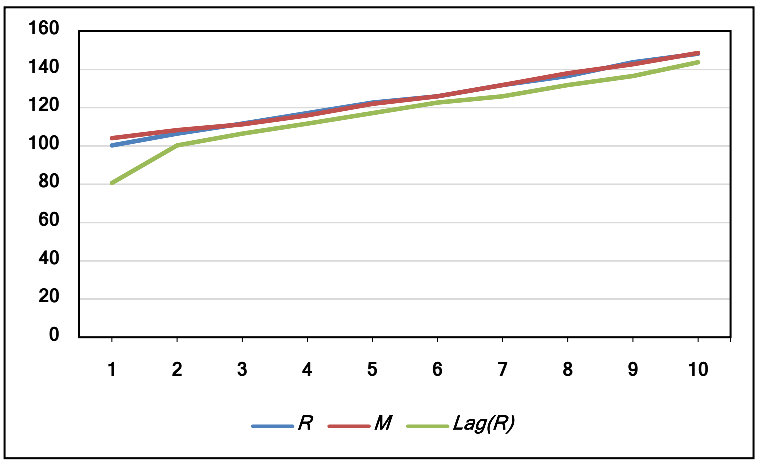

Figure 4. Revenue R, expenditure M, and lagged revenue $lag(R)$ in the experiment $q = 0.20$ (low q) (without RAND).

In summary, the study showed that OLSRT is, from the six methods assessed, the most efficient estimation method for *r*. In the same way, OLSU was proved to be the worst estimation method emphasizing the importance of forcing the regression through the origin. In fact, also LADRT, as forced through the origin, proved to be an efficient estimation method that was only slightly beaten by OLSRT. The incorporation of the revenue-expenditure restriction as adjustments in the OLS and ML estimates was efficient only in some exceptional cases. This restriction generally increased the dispersion of the estimates. In many experiments, RR was slightly more efficient than OLSU reflecting the effect of the bias (lambda). However, the experiments indicate that multicollinearity may statistically be of minor importance in the estimation of *r*. The contribution of the study is clearly in that it shows that the simplified Koyck transformation is very useful in estimating IRR when efficient estimators are employed. In spite of its simplicity, OLS as forced through the origin seems to provide us with accurate estimates of *r* in different, even very difficult circumstances. Experiments show that the unrestricted OLS (Laitinen & Laitinen, 2022) does not properly work in estimation and the incorporation of the revenue-expenditure restriction in OLS only makes the OLSRT less efficient (Laitinen & Laitinen, 2022; Laitinen, 2017).

The present study suffers from many limitations which can be relaxed in future studies. The present model was based on a simplification of the Koyck model and on a simplified model of a startup. In addition, only six estimation methods were compared with each other. Experiments were simple and based only two levels of the parameters. In future, the effect of the neglection of the MA part of the Koyck model should be analyzed more carefully. The model describing the development of a startup in its early years should be made more realistic to correspond better to the actual development. More estimation methods should be assessed in experiments to find the most efficient method. Experiments should be designed to conform better with the testing procedure. More parameters and levels should be applied in experiments. Statistical tests should be applied in comparing the efficiency of methods.

Conflicts of Interest

The authors declare no conflicts of interest regarding the publication of this paper.

References

Ak, B. K., Dechow, P. M., Sun, Y., & Wang, A. Y. (2013). The Use of Financial Ratio Models to Help Investors Predict and Interpret Significant Corporate Events. *Australian Journal of Management, 38,* 553-598. https://doi.org/10.1177/0312896213510714

Boumans, M., & Morgan, M. S. (2001). Ceteris Paribus Conditions: Materiality and the Application of Economic Theories. *Journal of Economic Methodology, 8,* 11-26. https://doi.org/10.1080/13501780010022794

Brief, R. P. (2013). *Estimating the Economic Rate of Return from Accounting Data (RLE Accounting)* (Vol. 16). Routledge. https://doi.org/10.4324/9781315886404

Chen, X., & Derezinski, M. (2021). Query Complexity of Least Absolute Deviation Regression via Robust Uniform Convergence. *Proceedings of Machine Learning Research, 134,* 1-36.

Davila, A., Foster, G., He, X., & Shimizu, C. (2015). The Rise and Fall of Startups: Creation and Destruction of Revenue and Jobs by Young Companies. *Australian Journal of Management, 40,* 6-35. https://doi.org/10.1177/0312896214525793

Duzan, H., & Shariff, N. S. B. M. (2015). Ridge Regression for Solving the Multicollinearity Problem: Review of Methods and Models. *Journal of Applied Sciences, 15,* 392-404. https://doi.org/10.3923/jas.2015.392.404

Eisenhauer, J. G. (2003). Regression through the Origin. *Teaching Statistics, 25,* 76-80. https://doi.org/10.1111/1467-9639.00136

Eurostat (2022). *Business Demography Statistics. Enterprise Survival Rate.* https://ec.europa.eu/eurostat/statistics-explained/index.php?title=Business_demography_statistics#Enterprise_survival_rate

Feenstra, D. W., & Wang, H. (2000). *Economic and Accounting Rates of Return.* University of Groningen. Research Institute SOM. SOM-Theme E Financial Markets and Institutions.

Fomby, T. B., Hill, C. R., & Johnson, S. R. (1984). *Advanced Econometric Methods.* Springer Verlag. https://doi.org/10.1007/978-1-4419-8746-4

Franses, P. H., & van Oest, R. (2004). *On the Econometrics of the Koyck Model.* Econometric Institute, Erasmus University Rotterdam.

Garnsey, E., Stam, E., & Heffernan, P. (2006). New Firm Growth: Exploring Processes and Paths. *Industry and Innovation, 13,* 1-20. https://doi.org/10.1080/13662710500513367

Hall, B. H. (2007). *Measuring the Returns to R&D: The Depreciation Problem.* Working Paper 13473. National Bureau of Economic Research. https://doi.org/10.3386/w13473

Hoerl, A. E., Kennard, R. W., & Baldwin, K. F. (1975). Ridge Regression Some Simulations. *Communications in Statistics, 4,* 105-123. https://doi.org/10.1080/03610927508827232

Huynh, K. P., Petrunia, R. J., & Voia, M. (2012). Duration of New Firms: The Role of Startup Financial Conditions, Industry and Aggregate Factors. *Structural Change and Economic Dynamics, 23,* 354-362. https://doi.org/10.1016/j.strueco.2012.03.008

Johnston, J. (1972). *Econometric Methods* (2nd ed., pp. 157-158). McGraw-Hill.

Klein, L. R. (1957). *The Estimation of Distributed Lags.* Cowles Foundation Discussion Papers, 251. https://elischolar.library.yale.edu/cowles-discussion-paper-series/251

Koyck, L. M. (1954). *Distributed Lags and Investment Analysis.* North-Holland Publishing Company.

Laitinen, E. K. (1992). Prediction of Failure of a Newly Founded Firm. *Journal of Business Venturing, 11,* 323-340. https://doi.org/10.1016/0883-9026(92)90005-C

Laitinen, E. K. (2017). Profitability Ratios in the Early Stages of a Startup. *The Journal of Entrepreneurial Finance, 19,* 1-28. https://doi.org/10.57229/2373-1761.1304

Laitinen, E. K., & Laitinen, T. (2022). Timing of Revenues and Expenses: Evidence from Finland. *Theoretical Economics Letters, 12,* 712-741. https://doi.org/10.4236/tel.2022.123040

Luckett, P. F. (1984). ARR vs. IRR: A Review and an Analysis. *Journal of Business Finance & Accounting, 11,* 213-231. https://doi.org/10.1111/j.1468-5957.1984.tb00070.x

Lukman, A. F., & Ayinde, K. (2016). Review and Classifications of the Ridge Parameter Estimation Techniques. *Hacettepe Journal of Mathematics and Statistics, 46,* 1-26. https://doi.org/10.15672/HJMS.201815671

Thanoon, F. H. (2015). Robust Regression by Least Absolute Deviations Method. *International Journal of Statistics and Applications, 5,* 109-112.

U.S. Bureau of Labor Statistics (2022). *Survival of Private Sector Establishments by Opening Year.* https://www.bls.gov/bdm/us_age_naics_00_table7.txt

Zingales, L. (1998). Survival of the Fittest or the Fattest? Exit and Financing in the Trucking Industry. *Journal of Finance, 53,* 905-938. https://doi.org/10.1111/0022-1082.00039